



PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,870

DATE: 02/25/2002

TIME: 11:45:48

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

Does Not Comply
Corrected Diskette Needed

6 <110> APPLICANT: Hartmann, Marcus
 7 Vohle, Peter
 8 Tiedke, Arno
 9 Baumert, Uwe
 11 <120> TITLE OF INVENTION: a-Hexosaminidase and a DNA Sequence Coding it Obtained
 12 from Ciliates and Use thereof
 14 <130> FILE REFERENCE: 012080us
 16 <140> CURRENT APPLICATION NUMBER: 09/914870
 17 <141> CURRENT FILING DATE: 2000-03-03
 19 <150> PRIOR APPLICATION NUMBER: DE19958979.8
 20 <151> PRIOR FILING DATE: 1999-12-08
 22 <150> PRIOR APPLICATION NUMBER: DE19909189.7
 23 <151> PRIOR FILING DATE: 1999-03-04
 25 <160> NUMBER OF SEQ ID NOS: 3
 27 <170> SOFTWARE: PatentIn Ver. 2.1

ERRORED SEQUENCES

29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 1656
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Tetrahymena
 34 <400> SEQUENCE: 1
 E--> 35 atgcaaaaga tacttttaaat tacttttcctt cttggaatag ctctcgtca aattactcct
 36 60
 E--> 37 ggcgttgacc ctatttcagc taagggttatg cctaaaccta agaattacac ttatggagat
 38 120
 E--> 39 ttgagcttac ttgtcactga tccttgcgga gtctcttaca gaccttctgt tgggtcagga
 40 180
 E--> 41 aaagtaccca accatgtcta tcaaattatt ggattctaca ctttgaatat tttcaattct
 42 240
 E--> 43 aacgaaaact cttgtgctat gtaaagagaa ttgtataaga atgaaacaac cattgaaaag
 44 300
 E--> 45 atgcgtagat tacaacattc ctaaaatata gtcttcgata tttttatcta agacgctgct
 46 360
 E--> 47 ttggccactg cagacacact cgaagacgaa tattatgatt tataaattta taataccaca
 48 420
 E--> 49 tattggaaat tgactgctaa caaatatggt ggtttactcc gtgggttaga aacttactct
 50 480
 E--> 51 caattattca cttaagacga agacactgaa gattgggtatt tgaataacat ccctatttct
 52 540
 E--> 53 attcaagatt aacctgacta catctacaga ggtcttatga tagattcagc cagacatttc

format error

60

120

↓

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/09/914,870

TIME: 11:45:48

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

```

54 600
E--> 55 ttatcagttg aaactatttt aaaaactatt gattctatgt tattcaacaa gttgaatgtt
56 660
E--> 57 ctccattggc acatcactga tactgaatcc ttccccttcc ctcttaaate attccctaatt
58 720
E--> 59 attactaaat atggagccta ctctaagaag aaacaataca gcttcgaaga catttaatac
60 780
E--> 61 attgtagact aagctctcaa caagggtatt taagttattc ctgaagtcga ttctccagga
62 840
E--> 63 cacgcttttt catgggctag atctccttaa ttctctagta ttgggtctatt atgtgattaa
64 900
E--> 65 tataatggat agtttagaccc aacactaaat ttaacttaca ctgctgttaa ggggtattatg
66 960
E--> 67 gaagatatga atacttaatt ctacactgct aagtatgttc attttggtgg tgatgaagtt
68 1020
E--> 69 gaagaataat gctggaataa acgccctgaa attaaggaat tcatgaatta aaataacatc
70 1080
E--> 71 tctacatata ctgatttgta gaattattac agaaagaact aagttaacat ttggaaatca
72 1140
E--> 73 attaattgcta ctaagcctgc tattttctgg gcagattcaa atactttgaa atatggctct
74 1200
E--> 75 gatgatatta ttcaatgggtg gggatctact catgattttt cttcaatcaa agatcttctt
76 1260
E--> 77 aacaaaataa ttttatcttt ctatgataat acttatttgg atgttggtga gggaaataga
78 1320
E--> 79 tatgggtggaa gttatggcag catgtataac tgggatgtct taaactcttt caatcctaga
80 1380
E--> 81 gttcctggaa ttaagggtga aattcttggg ggcgaaacat gcttatggag tgaaatgaat
82 1440
E--> 83 gatgattcta ctttaattcta aagactttgg acaagaaata gtgcatttgc tgaaagactt
84 1500
E--> 85 tggaacactg atgtgctaa caatgaaact tacaaaaacta gagctttagt tagcagaatg
86 1560
E--> 87 gtctttatgc aacaccgttt aactgctaga ggaatccctg cttctcctgt aacagttggg
88 1620
E--> 89 atttgtgaat aaaacctttc tctctgctac aattga
90 1656
206 <210> SEQ ID NO: 3
207 <211> LENGTH: 1837
208 <212> TYPE: DNA
209 <213> ORGANISM: Tetrahymena
211 <400> SEQUENCE: 3
E--> 212 cagcagtaat aaaaaattct aaatatattg attgtagcta tgcaaaagat acttttaatt
213 60
E--> 214 actttccttc ttggaatagc tctcgtcaa attactcctg gcgttgacct tatttcagct
215 120
E--> 216 aagggttatgc ctaaaccctaa gaattacact tatggagatt tgagcttact tgtcactgat
217 180
E--> 218 ccttgcgagg tctcttacag accttctgtt gggtcaggaa aagtacccaa ccatgtctat

```

*same**same*

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/09/914,870

TIME: 11:45:48

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

```

219 240
E--> 220 caaattattg gattctacac tttgaatatt ttcaattcta acgaaaactc ttgtgctatg
221 300
E--> 222 taaagagaat tgtataagaa tgaaacaacc attgaaaaga tgcgtagatt acaacattcc
223 360
E--> 224 taaaatatag tcttcgatat ttttatctaa gacgctgctt tggccactgc agacacactc
225 420
E--> 226 gaagacgaat attatgattt ataaatttat aataccacat attggaaatt gactgctaac
227 480
E--> 228 aaatatgttg gtttactccg tgggttagaa acttactctc aattattcac ttaagacgaa
229 540
E--> 230 gacactgaag attgggtattt gaataacatc cctattttcta ttcaagatta acctgactac
231 600
E--> 232 atctacagag gtcttatgat agattcagcc agacattttct tatcagttga aactatttta
233 660
E--> 234 aaaactattg attctatggt attcaacaag ttgaatgttc tccattggca catcactgat
235 720
E--> 236 actgaatcct tccccttccc tcttaaataa ttccctaata ttactaaata tggagcctac
237 780
E--> 238 tctaagaaga aacaatacag ctctgaagac atttaataca ttgtagacta agctctcaac
239 840
E--> 240 aagggtattt aagttattcc tgaagtcgat tctccaggac acgctttttc atgggctaga
241 900
E--> 242 tctccttaat tctctagtat tgggtctatta tgtgattaat ataatggata gttagacca
243 960
E--> 244 acactaaatt taacttacac tgctgttaag ggtattatgg aagatatgaa tacttaattc
245 1020
E--> 246 tacactgcta agtatgttca ttttgggtgg gatgaagttg aagaataatg ctggaataaa
247 1080
E--> 248 cgccctgaaa ttaaggaatt catgaattaa aataacatct ctacatatat tgattttag
249 1140
E--> 250 aattattaca gaaagaacta agttaacatt tggaaatcaa ttaatgctac taagcctgct
251 1200
E--> 252 attttctggg cagattcaaa tactttgaaa tatgggtcctg atgatattat tcaatgggtg
253 1260
E--> 254 ggatctactc atgatttttc ttcaatcaaa gatcttccta acaaaataat tttatctttc
255 1320
E--> 256 tatgataata cttatttgga tgttggtgag ggaaatagat atgggtggaag ttatggcagc
257 1380
E--> 258 atgtataact gggatgtctt aaactctttc aatcctagag ttcttggaat taagggtgaa
259 1440
E--> 260 attcttgggtg gcgaaacatg cttatggagt gaaatgaatg atgattctac ttaattctaa
261 1500
E--> 262 agactttgga caagaaatag tgcatttgct gaaagacttt ggaacactga tgctgctaac
263 1560
E--> 264 aatgaaactt acaaaaactag agcttttagtt agcagaatgg tctttatgca acaccgttta
265 1620
E--> 266 actgctagag gaatccctgc ttctcctgta acagttggta tttgtgaata aaacctttct
267 1680

```

same

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/09/914,870

TIME: 11:45:48

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

E--> 268 ctctgctaca attgattcta aatataaara tttaaataaat attttaagaa atattttttaa
269 1740 *same*

E--> 270 gaatatatttta gtataaaaaac tgtatttttaa ttgataaaaaa aaatataaat attattatta
271 1800

E--> 272 attgaattttt agctaaaaaaa aaaaaaaaaa aaaaaaaa
273 1837

E--> 276 - 5 - *delete*

VERIFICATION SUMMARY

DATE: 02/25/2002

PATENT APPLICATION: US/09/914,870

TIME: 11:45:49

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

L:35 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:60 SEQ:1

M:254 Repeated in SeqNo=1

L:212 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:60 SEQ:3

M:254 Repeated in SeqNo=3

STATISTICS SUMMARY

PATENT APPLICATION: US/09/914,870

DATE: 02/25/2002

TIME: 11:45:49

Input Set : A:\EP.txt

Output Set: N:\CRF3\02252002\I914870.raw

Application Serial Number: US/09/914,870

Alpha or Numeric: Numeric

Application Class:

Application File Date: 03-03-2000

Art Unit: PCT09

Software Application: PatentIN2.1

Total Number of Sequences: 3

Total Nucleotides: 3493

Total Amino Acids: 549

Number of Errors: 60

Number of Warnings: 0

Number of Corrections: 0

MESSAGE SUMMARY

254 E: 60 (No. of Bases conflict)